7 Sinnot JD. *Sex roles and aging: theory and research from a systems perspective.* London: Karger, 1986.
8 Annandale E, Hunt K. Masculinity, femininity and sex: an exploration of their relative contribution to explaining gender differences in health. *Sociol Health Illness* 1990;12:24-46.
9 Macintyre S, Hunt K, Sweeting H. Gender differences in health: are things really as simple as they seem? *Soc Sci Med* 1996;42:617-24.
10 King D. Gender blending. Medical perspectives and technology. In: Ekins R, King D, eds. *Blending genders.* London: Routledge, 1996:79-98.
11 Allison KR, Adlaf EM. Age and sex differences in physical inactivity among Ontario teenagers. *Can J Public Health* 1997;88:177-80.
12 Moynihan C, Bliss J, Davidson J, Burchell L, Horwich A. Evaluation of adjuvant psychological therapy in patients with testicular cancer: randomised controlled trial. *BMJ* 1998;316:429-35.
13 Kaplan M, Marks G. Appraisal of health risks: the roles of masculinity, femininity and sex. *Sociol Health Illness* 1995;17:206-21.
14 Hearn J. Research in men and masculinities: some sociological issues and possibilities. *Aust N Z J Sociol* 1994;30:47-70.
15 Bem S. The measurement of psychological androgyny. *J Clin Psychol* 1974;42):155-62.
16 Mitchel C. Relationship of femininity, masculinity and gender to attribution of responsibility. *Sex Roles* 1987;16:151-63.
17 Korzenny F. AIDS communication, beliefs and behaviours. Symposium on science communication: environmental and health research, University of Southern California, Los Angeles, 1988.
18 Stoltenberg J. *Refusing to be a man: essays on sex and justice.* Portland, OR: Breitenbush, 1989.
19 Moynihan C. Testicular cancer: the psychosocial problems of patients and their relatives. *Cancer Surv* 1987;6:477-510.
20 Seidler VJ. Men, heterosexualities and emotional life. In: Pile S, Thrift N, eds. *Mapping the subject: geographies of cultural transformation.* London: Routledge, 1995:170-91.
21 Seidler VJ. *Unreasonable men.* London, New York: Routledge, 1994.
22 Docherty T. *Postmodernism: a reader.* London: Harvester Wheatsheaf, 1993.
23 Ramirez A, Graham J, Richards M. Mental health of hospital consultants: the effects of stress and satisfaction of work. *Lancet* 1996;347:724-8.

Understanding controlled trials

What outcomes should be measured?

Martin Roland, David Torgerson

Many types of clinical, patient related, and economic outcomes can be measured in trials. The choice of one or more outcomes will depend on the nature of the study and the question it is trying to answer. Objectives can relate to different levels of observation and analysis, from the individual to the family, the community, and society as a whole.

If a trial is "explanatory"[1] then a single main measure of clinical outcome may be appropriate. For example, if a trial is designed to determine which of two antihypertensive agents is more effective at lowering blood pressure then hypertensive control will be the main outcome. Traditionally, clinical trials have used physiological or biomedical outcomes, but these may not be well related to clinical outcomes. One example of a surrogate outcome measure which misled investigators was the CD4 count in AIDS trials: this turned out to be a poor predictor of survival.[2] Thus the use of physiological surrogates which are not clearly related to health outcomes must be viewed with caution.

A range of health status measures have been developed to address the poor relation which may exist between clinical outcomes and outcomes that are important to patients. These attempt to capture the patient's experience using valid and reliable quantitative scales.[3 4] They generally aim to quantify the extent to which an illness affects a patient's ability to carry out a range of normal activities. They may be related to abilities across a wide range of activities or targeted at problems associated with specific diseases. A common approach is to use both a general and a disease specific measure within one trial.

In pragmatic trials a single outcome measure may be inadequate for clinicians and other healthcare decision makers to weigh up the risks, costs, and benefits of a given intervention. Several outcome measures are therefore commonly included. For example, in trials of back pain, the Cochrane Collaboration recommends that outcomes should include pain, functional status, ability to work, and satisfaction with treatment.[5] In another example a recent trial sought to compare evening and night care given by doctors from commercial deputising services with that given by a doctor from the patient's own practice; the outcomes included whether the patient was actually visited, what prescriptions were given, whether there was any difference in health outcome for patients, and whether care from one type of doctor was more likely to increase subsequent use of health services.[6 7] The use of a wide range of outcomes is likely to be more informative for decision makers than a single outcome measure.

The impact of a disease may extend beyond the individual to the family or carers—for example, in dementia[8]—so the outcomes measured might need to be extended to a wider group. Similarly, at a societal level, if an aim of the study is to influence resource allocation between different types of treatment then economic outcomes will need to be included.[9]

Although it is often advisable to use several different outcome measures, some have advocated that the very large trials needed to answer certain types of clinical problem should focus on a small number of very simple outcomes.[10] There is also a statistical drawback to using multiple outcome measures. Increasing the number of measures in a trial increases the probability that one will reach statistical significance on the basis of chance alone. When a research question requires that several separate outcomes should be separately assessed, this needs to be taken into account in the sample size calculation. In general, more subjects are needed when several outcomes are being measured.

National Primary Care Research and Development Centre, University of Manchester, Manchester M13 6PL
Martin Roland, *director of research and development*
Centre for Health Economics, University of York, York YO1 5DD
David Torgerson, *senior research fellow*
Correspondence to Dr Roland

BMJ 1998;317:1075

1 Roland MO, Torgerson DJ. What is a pragmatic trial? *BMJ* 1998;316;285.
2 Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;125:605-13.
3 Wilkin D, Hallam L, Doggett MA. *Measures of need and outcome for primary health care.* Oxford: Oxford University Press, 1992.
4 Bowling A. *Measuring disease: a review of disease specific quality of life measurement scales.* Buckingham: Open University Press, 1994.
5 Van Tulder MW, Assendelft WW, Koes BW, Bouter LM and the Cochrane Back Pain Editorial Board. Method guidelines for systematic reviews in the Cochrane Back Review Group for Spinal Disorders. *Spine* 1997;20:2323-30.
6 Cragg D, McKinley R, Roland M, Campbell S,Van F, Hastings A, et al. Comparison of out of hours care provided by patients' own general practitioners and commercial deputising services: a randomised controlled trial. 1. The process of care. *BMJ* 1997;314:187-9.
7 McKinley R, Cragg D, Campbell S, Hastings A, Roland M, French D, et al. Comparison of out of hours care provided by patients' own general practitioners and commercial deputising services: a randomised controlled trial. 2. The outcome of care. *BMJ* 1997;314:190-2.
8 O'Connor DW, Pollitt PA, Roth M, Brook CP, Reiss BB. Problems reported by relatives in a community study of dementia. *BMJ* 1990;156:835-41.
9 Torgerson DJ, Raftery J. Measuring outcomes in economic evaluations. *BMJ* 1998 (in press).
10 Peto R, Collins R, Gray R. Large scale randomised evidence: large, simple trials and overviews of trials. *J Clin Epidemiol* 1995;48:23-40.